

REMARKS

In the Office Action of October 23, 2003, the Examiner rejected claims 1-3, 5-8, 10-15, 17-22, 24-27, 29-32, 34, and 36-41 under 35 U.S.C. § 103(a) as being unpatentable over U.S. Patent No. 5,778,361 to Nanjo et al. ("Nanjo") in view of U.S. Patent No. 6,012,053 to Pant et al. ("Pant"); rejected claims 4, 9, 16, 23, 28, and 33 under 35 U.S.C. § 103(a) as being unpatentable over Nanjo in view of Pant, and further in view in view of U.S. Patent No. 6,385,602 to Tso et al. ("Tso"); and rejected claim 35 under 35 U.S.C. § 103(a) as being unpatentable over Nanjo in view of Pant, and further in view in view of U.S. Patent No. 6,134,554 to Freimann et al ("Friemann").

Because each claim has at least one feature not taught or suggested by the above-mentioned references, it is respectfully submitted that the claims are patentable over these references.

**A. NANJO AND PANT FAIL TO DISCLOSE OR SUGGEST AT LEAST ONE LIMITATION RECITED IN CLAIMS 1-3, 5-8, 10-15, 17-22, 24-27, 29-32, 34, AND 36-41, AND THEREFORE FAIL TO MAKE OBVIOUS THE INVENTION AS EMBODIED IN THOSE CLAIMS**

"To establish a *prima facie* case of obviousness...the prior art reference (or references when combined) must teach or suggest ***all*** the claim limitations. See M.P.E.P. § 706.02(j). Because claims 1-3, 5-8, 10-15, 17-22, 24-27, 29-32, 34, and 36-41 each contain the feature of selecting a semantic unit, which feature is neither disclosed or suggested by Nanjo and Pant, these references fail to make obvious the invention as defined by these claims.

Claim 1, for example, recites a method that includes selecting semantic units within a search query. The term "semantic unit," as defined by the pending application, refers to multiple terms that are considered to function as a "compound" that forms a single semantically meaningful unit. (Spec., page 2). Semantic units are selected in claim 1 through a method that includes identifying documents relating to a search query by matching individual search terms in the query to an index of a corpus and generating multiword substrings of the query in which each of the substrings includes at least two words. For each of the generated substrings, a value is calculated that corresponds to a comparison between one or more of the identified documents and the generated substring. Semantic units are selected from the generated multiword substrings based on the calculated values. Applicants submit that claim 1 is not disclosed or suggested by the combination of Nanjo and Pant.

Nanjo, in contrast to the invention recited in claim 1, does not disclose or suggest the identification of semantic units, much less the identification of semantic units from within a search query. The Examiner has previously conceded that Nanjo does not disclose identifying semantic units. (See, for example, Office Action of May 8, 2003, page 4, line 3). In the current Office Action, however, the Examiner now appears to contend that Nanjo's disclosure relating to "text in compound-word languages" and "combination of text strings" is an identification of semantic units. (Office Action, page 2).

Nanjo is directed to the indexing and searching of text in compound-word languages. Nanjo discloses, among other things, the creation of a content index

from a number of input documents. (Nanjo, col. 5, lines 22-42). In response to a user search query, the content index may be searched to obtain search results. (Id.). According to Nanjo, the content index may include tokens in which a single string may yield multiple tokens. Nanjo discusses this feature in detail at column 6, lines 22-21:

[T]he index is created by taking the collection of symbols forming the kanji character string, and creating a number of index terms each of a length the same as the step size, beginning with the first term in the string, and extending to the end of the kanji string, and thereafter progressively reducing the step size such that the last character in the kanji string is the last index term. In this manner, all kanji terms are taken in "chunks" of the step size or less, always beginning with one of the kanji symbols and always ending with a symbol at the end of a string of four or ending with the last symbol in the string.

The reason for step indexing is to cause the system to treat every kanji symbol or character as the potential beginning of a word. Furthermore, a step size is utilized that is equal to or longer than most words in the language in question. For Japanese, a step size of four is believed to be optimal. The document is then indexed by all tokens produced by the step indexing method. For example, the string "abcdefg" yields the tokens "abcd" "bcde", "cdef", "defg", "efg", "fg", and "g".

Applicants submit that the tokens created using the step indexing method of Nanjo are not semantic units. In particular, Nanjo appears to simply take multi-character sub-strings of a larger string. Nothing in Nanjo discloses or suggests that these sub-strings correspond to multi-word substrings that form a single semantic unit (i.e., a semantically meaningful unit).

The Examiner points to column 5, lines 23-28 and column 8, lines 21-33 as disclosing semantic units. Column 5, lines 23-28 of Nanjo state that the invention of Nanjo is particularly applicable in "compound-word languages such

as Japanese or Chinese” and generally discusses searching and indexing the context-index. This section of Nanjo does not disclose or suggest a semantic unit. The cited portion of column 8 similarly fails to disclose or suggest semantic units. More specifically, at column 8, lines 21-33, Nanjo discloses that users may enter search queries as a “combination of text strings” using, for example, Boolean operators. However, receiving a multi-word search query entered by a user is not analogous to, nor does it even suggest, identifying semantic units from a search query as multiword substrings from a search query. Although Nanjo may allow for multi-word search queries, Nanjo is completely unconcerned with, and makes no attempt to select, semantic units.

In view of the above discussion, Applicants submit that Nanjo fails to disclose or suggest at least one of the features recited in claim 1. Nanjo, for instance, does not select semantic units within a search query. Thus, it is understandable that Nanjo further fails to suggest or even mention at least one of the recited features of selecting a semantic unit by identifying documents relating to a search query by matching individual search terms in the query to an index of a corpus and generating multiword substrings of the query in which each of the substrings includes at least two words. It is further understandable that Nanjo does not suggest or even mention the feature of calculating values and select semantic units from the generated multiword substrings based on the calculated values, as recited in claim 1.

Applicants submit that Pant fails to cure the above-noted deficiencies of Nanjo. Pant discloses a system for performing searches on a collection of

information through which results from a search query are ranked according to user specified relevance factors that allow the user to control how the search results are presented. Column 7, lines 7-50, of Pant, discusses a number of relevance factors through which documents that contain at least some terms of a user search query can be ranked. For example, Pant states that "if a user enters a query that has six search terms, than [sic.] documents which contain all six search terms are generally considered more relevant than documents which contain only five of the six search terms." (Pant, col. 7, lines 7-12). Another relevance factor mentioned by Pant "is the ordering of search terms in the document. That is, if the query terms appear in their given order in a document, than [sic.] a relevance bonus may be applied to the document." (Pant, col. 7, lines 14-18). Column 10, lines 38-49 of Pant is similarly directed to weighting a document for relevance based on the number of search terms that occur in the document.

Thus, the sections of Pant cited by the Examiner are merely related to relevance factors that are used to rank documents, and not the calculation of values that are then used to select semantic units, for example, as recited in claim 1. Although Pant may generally calculate "values" based on the terms in a document and search queries, this does not even suggest calculation of values that are then used to select semantic units as recited in claim 1.

For at least the reasons discussed above, Applicants submit that Nanjo and Pant, even if combined as the Examiner suggests, still fail to disclose or suggest at least one of the features recited in claim 1, and thus, fail even in

combination to make obvious the invention embodied by claim 1. Thus, it is respectfully submitted that the rejection of claim 1 should be withdrawn. At least by virtue of their dependency on claim 1, the rejection of claims 2, 3, 5, and 37 should also be withdrawn.

Independent claim 6 recites a number of features similar to those recited in claim 1, including "generating a plurality of multiword substrings of the query" and "selecting semantic units from the generated multiword substrings based on the calculated values." For reasons similar to those given above, Applicants submit that Nanjo and Pant, either taken alone or in combination, do not disclose or suggest these features of claim 6. Accordingly, the rejection of this claim should thus be withdrawn. At least by virtue of their dependency from claim 6, the rejection of claims 7, 8, 10, and 38 should also be withdrawn.

Independent claim 11 is directed to a system and also includes features similar to those recited in claim 1. Thus, for reasons similar to those given with respect to claim 1, the rejection of claim 11 should also be withdrawn. The rejection of claims 12-15, 17, and 39, which depend from claim 11, should also be withdrawn, at least by virtue of their dependency.

Independent claim 18 recites, inter alia, the feature of a ranking component configured to return a list of documents ordered by relevance in response to a search query and a semantic unit component configured to locate semantic units, having a plurality of words, in search queries entered by a user based on a predetermined number of most relevant documents in the list of

documents returned by the ranking component. As previously discussed, neither Nanjo nor Pant discloses or suggests locating semantic units in search queries, much less locating semantic units based on a predetermined number of most relevant documents in the list of documents returned by the ranking component. Accordingly, the rejection of claim 18 should be withdrawn for at least one of the above-mentioned reasons discussed with respect to claim 1.

Independent claim 25 recites features similar to those recited in claim 1. Independent claims 30 and 36 recite features similar to those recited in claim 6. Thus, for reasons similar to those given above regarding claims 1 and 6, the rejection of these claims should also be withdrawn. As such, the rejection of claims 26, 27, 29, 31, 32, 34, 40, and 41 at least by virtue of their dependency from one of claims 25 or 30, should also be withdrawn.

**B. *BECAUSE THE CITED REFERENCES FAIL TO DISCLOSE OR SUGGEST AT LEAST ONE LIMITATION RECITED IN THE INDEPENDENT CLAIMS, AS DISCUSSED ABOVE, BY VIRTUE OF THEIR DEPENDENCY, CLAIMS 4, 9, 16, 23, 28, and 33 ARE ALSO NON-OBVIOUS AND PATENTABLE OVER THE CITED REFERENCES***

Dependent claims 4, 9, 16, 23, 28, and 33 stand rejected under 35 U.S.C. §103(a) as being unpatentable over Nanjo in view of Pant, and further in view of Tso. Because each of the claims incorporate, by virtue of dependency, at least one limitation not suggested or even mentioned by the cited references (see arguments above), the cited references fail to make obvious these claims at least for this reason.

this application. Applicants therefore request the Examiner's reconsideration and reexamination of the application, and the timely allowance of the pending claims. If the Examiner believes that the application is still not in condition for allowance, Applicants invite the Examiner to call the undersigned at the below-listed number to discuss any further issues that the Examiner may have.

To the extent necessary, a petition for an extension of time under 37 C.F.R. § 1.136 is hereby made. Please charge any shortage in fees due in connection with the filing of this paper, including extension of time fees, to Deposit Account 50-1070 and please credit any excess fees to such deposit account.

Respectfully submitted,

HARRITY & SNYDER, L.L.P.

By:   
Brian E Ledell  
Reg. No. 42,784

Date: March 23, 2004

11240 Waples Mill Road  
Suite 300  
Fairfax, Virginia 22030  
Telephone: 571-432-0800  
Facsimile: 571-432-0808  
Customer Number: 26615